TRANSLATION FROM FRENCH

E-620E
c.2

ANALYSIS OF LANGUAGE MOBILITY: INDEXES, OBSERVATIONS AND MODELS

by

Réjean Lachapelle*

Research Paper No. 12b, May, 1985
Social and Economic Studies Division
Statistics Canada

Article submitted to Cahiers québécois de démographie.

December 1984

* Wherever possible, we have taken into account the comments of Leroy O. Stone, which we gratefully acknowledge. We also wish to thank Luc Albert and Luc St-Amour for their technical assistance and Audrey Sirois for typing the various drafts of this article. The views expressed in this paper are the responsibility of the author alone, not of Statistics Canada.

Summary

Two methods of analysing language mobility are presented. The first is based on the distinction between structural mobility, which produces a change in the final (or home-language) composition relative to the initial (or mother-tongue) composition, and exchange mobility, which measures the frequency of balanced transfers between each pair of languages. The two corresponding sets of parameters, together with mother-tongue composition, summarize all the information contained in a cross-tabulation of the population by mother tongue and home language. All other indexes can be defined using these coefficients. This approach resolves the paradoxes perceived in the findings of various studies of language mobility. The second method is more conventional in that it relies on the theory of stationary Markov chains. It entails the transformation of language mobility rates into an equilibrium composition. The latter is then compared with the observed compositions. Except in very special cases, this method yields less useful results than the first one, which is better suited to the concerns underlying language mobility research.

Introduction

Since the release of 1971 Census data on mother tongue[1] and home language,[2] there has been a great deal of research into language mobility and the events it generates, language transfers (see especially Castonguay, 1983a; De Vries and Vallee, 1980; Gryz, 1980; Lachapelle and Henripin, 1982). A number of analytical approaches were developed and tested, and many different indexes were proposed and applied to explore the various facets of the subject. While this diversity of approaches and methods firmly established a number of findings, it also led to some apparently contradictory conclusions (Castonguay, 1980; Lachapelle, 1980).

There are two possible approaches to the study of language mobility: either one is interested in its impact on structures or more general phenomena, or one attempts to uncover and measure the processes that cause it or the factors that account for its variations. In the first case, language mobility is an explanatory variable or, in statistical jargon, an independent variable. The focus is on the effects it has on linguistic composition (Castonguay, 1983b; Castonguay and Marion, 1974; Lachapelle and Henripin, 1982). In the second case, language mobility is seen as a dependent or explained variable. Some researchers have taken a conventional approach (De Vries and Vallee, 1980; Gryz, 1980), sometimes referred to as factor analysis. It involves examining how the phenomenon being studied is affected by different variables (sex, age, marital status, place of birth, level of schooling, place of residence and so on), taken individually or jointly. This approach has the advantage of not requiring any assumptions. Because of its exploratory nature, it is well-suited to the study of poorly understood phenomena. Others have opted to skip this step and employ the language of processes from the outset. They have shown that it is productive to regard language mobility as the outcome of a series of independent subprocesses associated with the life cycle: acquisition of a second language, choice of spouse and language orientation in homolinguistic and heterolinguistic couples (Castonguay, 1983a; Lachapelle, 1983). While this avenue often leads to

conclusions, it is predicated on a large number of assumptions, which, though plausible, are difficult to test directly because of the limitations of census data.

All these approaches and methods have a common starting point - a matrix or table containing the joint mother tongue-home language distribution of a given subpopulation. From it are derived ad hoc indexes that summarize part of the information in the table. These indexes are discussed in the first section of this paper. Additional indexes result from a comprehensive analysis or are defined in relation to a model. The modelling process is aimed at reconciling the apparently contradictory statements arising from an examination of various ad hoc indexes. It also makes the analysis more coherent. In some cases, the model suggests new indexes; in others, it lends new meaning to an index in common use. These topics will be covered in the second section. The proposed measures and indexes will be illustrated with Canadian census data. Comments will be kept to a minimum since the objective here is not to explore a real situation but to build and calibrate analytical tools.

Ad Hoc Indexes

Let n_{ij} be the number of persons with initial language (mother tongue) i and final language (home language) j. The number of people with mother tongue i will be denoted $n_{i.}$ and the number with home language i $n_{.i}$. Many indicators can be produced using these data. Some of them are based exclusively on the table's marginal totals (the $n_{i.}$ and the $n_{.i}$), while others involve components from the body of the table. In both cases, it is possible to look at either differences in absolute frequencies or ratios, rates or propensities. We shall first examine those indicators that are, or are derived from, net differences.

Absolute Differences

Language groups can be divided into two classes depending on whether they gain or lose as a result of linguistic mobility. We therefore need to measure the net transfers experienced by group i, as follows:

$$TN_i = n_{.i} - n_{i.} \qquad (I)$$

The sum of these quantities must be zero, since

$$n_{..} = \sum_i n_{.i} = \sum_i n_{i.} \qquad (II)$$

If no group is to gain from this process, the net transfers must all be zero or, in other words, the population's home-language composition must be the same as its mother-tongue composition.

We shall consider the languages in pairs. Net transfers to language i from language j are denoted TN_{ij} ($= n_{ji} - n_{ij}$). We thus obtain a decomposition of the net transfers of group i, since

$$TN_i = \sum_{j \neq i} TN_{ij} \qquad (III)$$

The TN_{ij} also provide an indication of the dominance ranking of the languages concerned. If we have $TN_{ij} > 0$, $TN_{jk} > 0$ and $TN_{ik} > 0$, it follows not only that language i dominates language j and the latter prevails over language k, but also that the dominance relation is transitive (language i prevails over language k). Moreover, when $TN_{ik} > TN_{jk}$, the relation may be described as strongly transitive.

We shall illustrate these considerations with data taken from the 1981 Census of Canada (Table 1). On the basis of net language transfers, English prevails over French in all parts of the country and both official languages dominate the "other" languages. English is a dominant

TABLE 1
Net Language Transfers, Canada, Quebec and Canada Excluding Quebec, 1981[a]

(In thousands[b])

Area and Language	Net Transfers to		
	English	French	Other Languages
Canada			
English	-	-288	-1,387
French	288	-	-35
Other	1,387	35	-
Total	1,675	-253	-1,422
Quebec			
English	-	-24	-90
French	24	-	-33
Other	90	33	-
Total	114	8	-123
Canada excl. Quebec			
English	-	-264	-1,297
French	264	-	-3
Other	1,297	3	-
Total	1,561	-262	-1,300

a Excluding inmates.
b Because of rounding, components do not always add up to the totals.

Source: Statistics Canada, 1981 Census of Canada; calculations by the author.

language and French a subdominant one; the "other" languages are dominated. As expected, the dominance relation is strongly transitive. These statements apply both to Quebec and to the rest of the country, which at least at first glance seems surprising.

Let us analyse the problem more closely. The language transfers between a dominant language (i) and a dominated one (k) are highly asymmetric, n_{ik} being negligible in comparison with n_{ki}. This can be clearly demonstrated by measuring the efficiency of language transfers, as follows:

$$EF_{ik} = \max (n_{ik}, n_{ki}) / (n_{ik} + n_{ki}) \qquad (IV)$$

This index takes values ranging between 0.5 and 1. When the language transfers between the two groups are perfectly balanced, the efficiency index is equal to 0.5. It is unity when there are no transfers from the dominant group to the dominated group. As a first approximation, transfers may be considered inefficient when the index is below 0.75 and efficient when it exceeds this level.

In Quebec, as we have seen, English prevails over French and both languages clearly dominate the "other" languages. This conclusion is substantiated by the efficiency indexes (Table 2). Transfers between English and French are the least efficient (0.6); by contrast, transfers between English and "other" languages are highly efficient (0.9). As expected, the efficiency of transfers between French and "other" languages falls somewhere in the middle (0.8). In the rest of Canada, however, there is little difference between the efficiency of English-French transfers (almost 0.9) and English-"other" transfers (over 0.9). This suggests that French is a half dominated language outside Quebec, a conjecture supported by the low efficiency index of French-"other" transfers (0.7).

TABLE 2
Net Transfer Efficiency Index, Canada, Quebec and Canada Excluding Québec, 1981[a]

(In percent)

Area	Transfers Between		
	English and French	English and Other Languages	French and Other Languages
Canada	77	93	76
Quebec	56	90	77
Canada excl. Québec	88	93	66

[a] Excluding inmates.

Source: Statistics Canada, 1981 Census of Canada; calculations by the author.

This conclusion is substantiated by other indexes, such as the proportion of the "other" group's net losses that goes to the English group (the complement of this proportion is the French group's share), which is 73 per cent in Quebec and more than 99 per cent elsewhere in Canada. We obtain similar results for the proportion of gross transfers from "other" languages to English (69 per cent in Quebec and over 99 per cent in the rest of Canada).

Thus, by examining a few simple indexes, we have been able to break down the languages spoken in Canada's two major language areas into four classes: dominant language (English everywhere), subdominant language (French in Quebec), half dominated language (French elsewhere in Canada) and dominated language (other languages everywhere). Is this classification an underlying structure, or an artifact of the indexes used?

Ratios

For the purposes of comparison, it is frequently preferable to examine relative rather than absolute frequencies. Let P_{ij} be the proportion of the population with mother tongue i and home language j. It follows immediately that

$$P_{i.} = \sum_j P_{ij}$$

$$P_{.i} = \sum_j P_{ji}$$

$$\sum_i P_{i.} = \sum_i P_{.i} = 1 \qquad \text{(V)}$$

where $P_{i.}$ and $P_{.i}$ denote the proportions of the population with i as mother tongue and home language respectively. Clearly, the indexes described above can be reformulated in terms of relative frequencies. This has no effect on the sign or the meaning of the differences. The efficiency index also remains unchanged by this transformation.

The net transfer rate has often been used as a measure of the net effect of transfers on a language group (Rochon-Lesage and Maheu, 1974):

$$tn_i = (P_{.i} - P_{i.})/P_{i.} = (n_{.i} - n_{i.})/n_{i.}$$

$$= TN_i/n_{i.} \qquad \text{(VI)}$$

This rate is positive when a group gains through language transfers and negative when it loses. If one is interested only in language groups that lose, it seems natural to consider the opposite of this rate (Castonguay and Marion, 1974), the net linguistic mobility rate (tm_i). Another measure has also been used, the index of linguistic continuity (Lachapelle and Henripin, 1982), defined as follows:

$$c_i = n_{.i}/n_{i.} = P_{.i}/P_{i.} \qquad \text{(VII)}$$

The relations between these indexes are very simple:

$$tm_i = - tn_i$$

and $$tm_i = 1 - c_i \qquad \text{(VIII)}$$

They should therefore invariably lead to the same conclusions. This means that depending on what the analysis requires, one can choose any one of these quantities to measure the net effect of linguistic mobility. Yet the probablity of error inherent in each one must be taken into account.

Any index is, of course, only an approximation to the actual value it is intended to measure. Let e_i be the measurement error affecting tn_i. From the preceding equation, then, we have

$$tn_i + e_i = c_i + e_i - 1 \qquad \text{(IX)}$$

The error made in estimating the net transfer rate is equal in magnitude to the error affecting the continuity index. This same cannot be said, however, for the absolute value of the relative error (er). We have

$$er\ (tn_i) = \left| e_i / tn_i \right| = er\ (tm_i)$$

and

$$er\ (c_i) = \left| e_i / c_i \right| \qquad \text{(X)}$$

It follows directly that

$$er\ (c_i) \ /\ er(tm_i) = \left| tm_i \right| \ /\ c_i \qquad \text{(XI)}$$

The relative error affecting the net rates is equal to that affecting the continuity index only when $c_i = tm_i = -tn_i = 0.5$. When the continuity index is greater than 0.5, the relative error may be much higher for the net rates than for the continuity index. The converse is true when the continuity index is less than 0.5.

To make this clearer, let us take an example. Suppose that within a given population, the observed continuity index for a particular language group is 1.001. Since the actual value is 0.999, the measurement error

is -0.002. The relative error is therefore 2 per 1,000 (0.002/1.001). The relative error for the net language transfer rate, on the other hand, is much larger. The difference between its observed value (0.001) and its actual value (-0.001) is identical to the absolute error affecting the continuity index, but the relative error (0.002/0.001) is a thousand times larger.

Is it necessary, then, to select the measure with the smallest relative error in each particular case? Such a restriction would probably be excessive. It would also complicate the analysis unnecessarily since the measure selected might vary from one language group or population to another. Nevertheless, in interpreting discrepancies, the relations between the relative errors of the various measures must be kept in mind.

When a group gains from the process of language mobility, its net transfer rate is positive and its continuity index is greater than 1. In all parts of Canada, the English group's continuity index is higher than that of the French group (Table 3); the "other" group's index is invariably lower than that of the French group. In Quebec, the French group's index is slightly over 1, while in the rest of Canada it is far below that (0.7). These findings are consistent with the results produced by examining the differences between absolute numbers of language transfers.

Is the continuity index a measure of the relative power of attraction of the various languages involved? Such an interpretation is not without difficulties, since it is hard to see how one could say that English has a greater power of attraction in Quebec (1.16) than in the rest of Canada (1.11). It should be noted that the continuity index's range depends on the initial (or mother-tongue) linguistic composition; for a given language group, it must fall between 0 and $1/p_i$. Hence, the English group's continuity index has a maximum value of approximately 9 in Quebec compared with only 1.26 elsewhere in Canada, which makes interpretation of the observed difference very tricky indeed. Moreover, the proportions of net language transfers (Table 3) clearly show that language mobility results in larger gains for the English group in the rest of Canada than in Quebec.

TABLE 3
Ad Hoc Indexes Used to Describe the Language Situation, Canada, Quebec and Canada Excluding Quebec, 1981[a]

(In percent)

	Composition of the Population by		Net Transfer Rate	Net Linguistic Mobility Rate	Continuity Index	Relative Frequency of Net Language Transfers[d]
	Mother Tongue[b]	Home Language				
Canada						
English	61.3	68.2	11.4	-11.4	111.4	7.0
French	25.7	24.6	-4.1	4.1	95.9	-1.1
Other	13.0	7.2	-45.1	45.1	54.9	-5.9
(Total Number[c])	(24,343)	(24,083)				
Quebec						
English	11.0	12.7	16.4	-16.4	116.4	1.8
French	82.4	82.5	0.2	-0.2	100.2	0.1
Other	6.6	4.8	-28.8	28.8	71.2	-1.9
(Total Number[c])	(6,438)	(6,369)				
Canada Excluding Quebec	79.4	88.2	11.1	-11.1	111.1	8.8
English	5.3	3.8	-28.2	28.2	71.8	-1.5
French	15.4	8.1	-47.6	47.6	52.4	-7.3
Other	(17,905)	(17,714)				
(Total Number[c])						

a Excluding inmates.
b Inmates are included in this column only.
c In thousands.
d This index represents $P_{.i}-P_{i.}$ in the population being studies (excluding inmates).

Source: Statistics Canada, 1981 Census of Canada; calculations by the author.

It has been proposed that language mobility rates be used to measure the intensity of the phenomenon for a given language in various populations (Lachapelle and Henripin, 1982). Let t_{ij} be the rate of linguistic mobility from initial language i to final language j. The value of this rate is given by

$$t_{ij} = n_{ij}/n_{i.} = p_{ij}/p_{i.} \qquad \text{(XII)}$$

When i = j, this is said to be a rate of language retention. For any language i, it is always true that

$$\sum_j t_{ij} = 1 \qquad \text{(XIII)}$$

The retention rate and its complement (the general language mobility rate) are not unrelated to the quantities discussed earlier. It can readily be shown that

$$c_i \geq p_{ii}$$

and that

$$tm_i \leq 1-p_{ii} \qquad \text{(XIV)}$$

The continuity index is an overestimate of the retention rate, and the net linguistic mobility rate is an underestimate of the general linguistic mobility rate. However, they often provide satisfactory approximations for dominated languages that receive few transfers from other languages. In Canada, this statement applies primarily to the "other" group and the French group outside Quebec.

It has also been suggested that linguistic mobility rates be used to describe the relative status of the various languages in a given population (Kralt, 1977). While findings based on these new indexes do not conflict with earlier assessments of the situation in the territories and provinces other than Quebec, they seem to indicate that in Quebec, contrary to expectations, French dominates English (Table 4), as the rate of mobility from English to French (12 per cent) is much higher than the

rate from French to English (2 per cent). This paradox is by no means new, having been discovered in the 1971 Census data (Castonguay, 1980; Lachapelle, 1980). It merits closer study.

TABLE 4
Language Retention Rate and Language Mobility Rate, Canada, Quebec and Canada Excluding Quebec, 1981[a]

(In percent)

Area and Mother Tongue	Home Language		
	English	French	Other
Canada			
English	98.4	0.8	0.8
French	6.6	93.1	0.3
Other	47.4	1.6	51.0
Quebec			
English	86.5	11.8	1.7
French	2.0	97.7	0.3
Other	23.9	10.9	65.2
Canada excl. Quebec			
English	99.0	0.3	0.7
French	32.8	66.9	0.3
Other	51.1	0.2	48.7

[a] Excluding inmates.

Source: Statistics Canada, 1981 Census of Canada; calculations by the author.

Indexes Derived from Models

The number of empirical indexes that can be developed is limited only by researchers' imaginations. This surfeit of analytical tools is illusory, however, because the number of independent parameters is not unlimited. A cross-tabulation of n mother tongues and n home languages has n^2 degrees of freedom ($n^2 - 1$ if relative frequencies are used),

which means that it is impossible to define more than n^2 independent parameters. This is a general constraint. Needless to say, it does not imply that we can determine which are the best indexes. That depends on one's point of view.

For the study of language mobility, two approaches immediately come to mind. Either we concern ourselves primarily with its structural effects, paying special attention to the table's marginal totals (mother-tongue and home-language compositions), or we concentrate on propensities, chiefly by looking at mobility rates. We shall develop these two approaches separately and in detail. For convenience, we shall assume that there are only two languages involved (i and j).

From Linguistic Composition to Mobility Rates

Suppose we know the mother-tongue composition (the $n_{i.}$ and the $p_{i.}$) and the home-language composition (the $n_{.i}$ and the $p_{.i}$) but not the absolute and relative frequencies of transfers from i to j (n_{ij} and p_{ij}) and from j to i (n_{ji} and p_{ji}). While it is impossible under these conditions to determine the exact value of the linguistic mobility rates (t_{ij} and t_{ji}), the range of these rates is not independent of the observed linguistic compositions.

We may suppose, without loss of generality, that group i always gains through language exchanges, or at least never loses. This amounts to assuming that $c_i \geq 1 \geq c_j$ or equivalently that $p_{.i} \geq p_{i.}$. Since we know $p_{ji} - p_{ij} = p_{.i} - p_{i.}$, these assumptions imply that

$$p_{ji} \geq p_{ij}$$

and that

$$(1-p_{i.})\, t_{ji} \geq p_{i.}\, t_{ij} \qquad \text{(XV)}$$

The latter inequality clearly indicates that mobility rates do not vary independently of the initial linguistic composition.

To understand this more clearly, let us consider the situation in which the group that gains (i) also makes up the majority of the population ($P_{i.} \geq 1-P_{i.}$). We therefore have

$$t_{ji}/t_{ij} \geq P_{i.}/(1 - P_{i.}) \geq 1 \qquad \text{(XVI)}$$

The minority group always has a higher mobility rate than the majority group. There is one exception to this, but it is trivial: when $c_i = c_j = 1$, both language transfers and mobility rates may be zero. Moreover, if the language groups are equal in size, the rates of mobility - regardless of its strength - must be equal.

Because of constraints due to the linguistic compositions, the mobility rates do not vary between 0 and 1. It can easily be shown that the majority group's rate (t_{ij}) ranges between 0 and $P_{.j}/P_{i.}$, while the minority group's rate (t_{ji}) can only take values between tm_j and 1, where tm_j is, as mentioned above, the net mobility rate of group j.

It seems reasonable to standardize the mobility rates so that they have the same range, from 0 to 1. Using the letter g (g_{ij} and g_{ji}) to denote the standardized rates, we have

$$g_{ij} = t_{ij} \,/\, P_{.j}/P_{i.} = P_{ij}/P_{.j}$$

and

$$\begin{aligned} g_{ji} &= (t_{ji}-tm_j) \,/\, (1-tm_j) \\ &= (P_{ji} - (P_{.i}-P_{i.})) \,/(1-P_{.i}) \\ &= P_{ij}/P_{.j} = g_{ij} \end{aligned} \qquad \text{(XVII)}$$

The standardized rates are always equal to one another. Parameter g ($= P_{ij}/P_{.j}$) represents the language mobility between the two groups. Note that the mobility rates can be expressed in terms of this parameter:

$$t_{ij} = g\,(1 - P_{.i}) \,/\, P_{i.} \;=\; g(1-tm_j)\,(1-P_{i.})/P_{i.}$$

and

$$t_{ji} = tm_j + g\,(1-tm_j)$$

or

$$t_{ji} = tm_j + t_{ij}\, P_{i.} \,/\, (1 - P_{i.}) \qquad \text{(XVIII)}$$

When g is zero, the dominant majority group has a mobility rate of zero, and the minority group's rate declines to a minimum ($^{tm}_{j}$). Conversely, when g equals 1, both rates are at a maximum.

In addition to the terms denoting components of the linguistic compositions, two parameters appear in the foregoing equations, g and $^{tm}_{j}$. As we have seen, $^{tm}_{j}$ measures the transfer imbalance detrimental to group j. Can we interpret g as a measure of the frequency of balanced transfers? If $^{tm}_{j}$ is zero, we obviously have $P_{i.} = P_{.i}$ and hence

$$t_{ij} = g\ (1-P_{i.})/P_{.i}$$
$$t_{ji} = g \qquad \text{(XIX)}$$

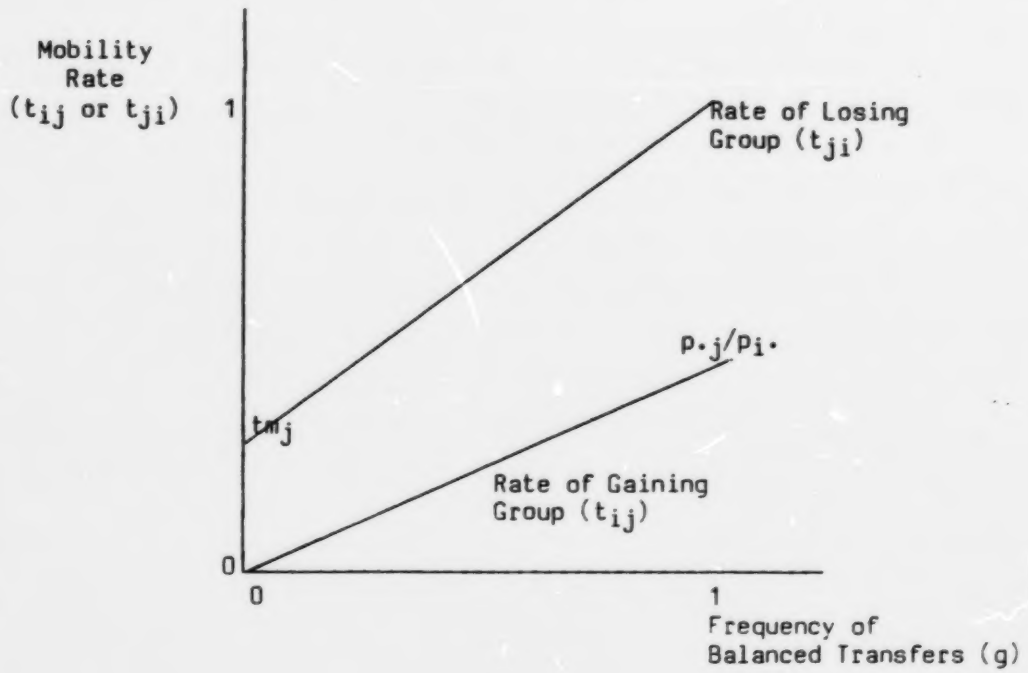
This implies that the language transfers are balanced, since

$$P_{ij} = P_{i.}\ t_{ij} = g\ (1-P_{i.})$$
and $$P_{ji} = (1-P_{i.})\ t_{ji} = g\ (1-P_{i.}) \qquad \text{(XX)}$$

Three independent parameters summarize all the information contained in the table: (1) the ratio of the gaining group's population to the total population ($P_{i.}$), which describes the mother-tongue composition; (2) the net mobility rate of the losing group ($^{tm}_{j}$), which measures the imbalance in language transfers; and (3) a coefficient (g) that represents the frequency of balanced exchanges between the two groups.

By virtue of (XVIII), there is a linear relation between the mobility rates and g (Figure 1). These lines are fully defined when the linguistic compositions, or equivalently parameters $P_{i.}$ and $^{tm}_{j}$, are known. The quantity g specifies the value of the mobility rates in the population being studied.

Figure 1. The Gaining Group is in the Majority



Is it possible to find analogous relations in the case where the dominant group is in the minority ($P_{i\cdot} \leq 1-P_{i\cdot}$)? We shall first attempt to determine the endpoints of the range of the mobility rates. Since group i is in the minority, its mobility rate (t_{ij}) may vary anywhere from 0 to 1. The majority group's mobility rate (t_{ji}), on the other hand, has a much narrower range, its lower limit being tm_j and its upper limit $P_{\cdot i}/(1-P_{i\cdot})$. To obtain these results, we must assume that $P_{j\cdot} > P_{\cdot i}$. This condition is not always satisfied, particularly when the dominant group makes up nearly half the population.[3]

What is the value of g? Through an argument similar to the one used above, it can easily be proved that

$$g_{ij} = t_{ij}$$

and $$g_{ji} = (t_{ji}-tm_j) \,/\, ((p_{.i}/(1-p_{i.})) - tm_j)$$

$$= (p_{ji} - (p_{.i}-p_{i.})) \,/\, p_{i.}$$

$$= p_{ij}/p_{i.} = g_{ij} \qquad \text{(XXI)}$$

Although g has the same numerator as in equation (XVII), the denominator is different. In fact, in all cases, g is the ratio of the smaller number of language transfers to the smallest marginal total:

$$g = \min (p_{ij},p_{ji}) \,/\, \min (p_{i.},p_{j.},p_{.i},p_{.j}) \qquad \text{(XXII)}$$

The complement of g is clearly a general index of maintenance or immobility (Boudon, 1973).

Now we must express the mobility rates in terms of g, using equation (XXI):

$$t_{ij} = g$$

and $$t_{ji} = tm_j + g\, p_{i.}/(1-p_{.i})$$

or $$t_{ji} = tm_j + t_{ij}\, p_{i.}/(1-p_{i.}) \qquad \text{(XXIII)}$$

The preceding equation is identical with the third equation in (XVIII). In that case, we had $t_{ji} > t_{ij}$ because the dominant group was in the majority. What happens when this group is in the minority? First let us define a such that $t_{ij} = t_{ji} = a$. From the third equation in (XXIII), we have

$$a - tm_j = a\, p_{i.}/(1-p_{i.})$$

$$a\,(1-2p_{i.})/(1-p_{i.}) = tm_j$$

$$a = (p_{.i}-p_{i.})/(1-2p_{i.}) \qquad \text{(XXIV)}$$

This coefficient is less than or equal to 1, provided that $p_{j.} \geq p_{.i}$. This condition cannot be satisfied when the dominant group is a very large minority. It can be shown that this situation is equivalent to the case in which the dominant group is in the majority. Note also that a is always larger than tm_j.

Under what conditions is it possible for $t_{ij} > t_{ji}$? This can also be written

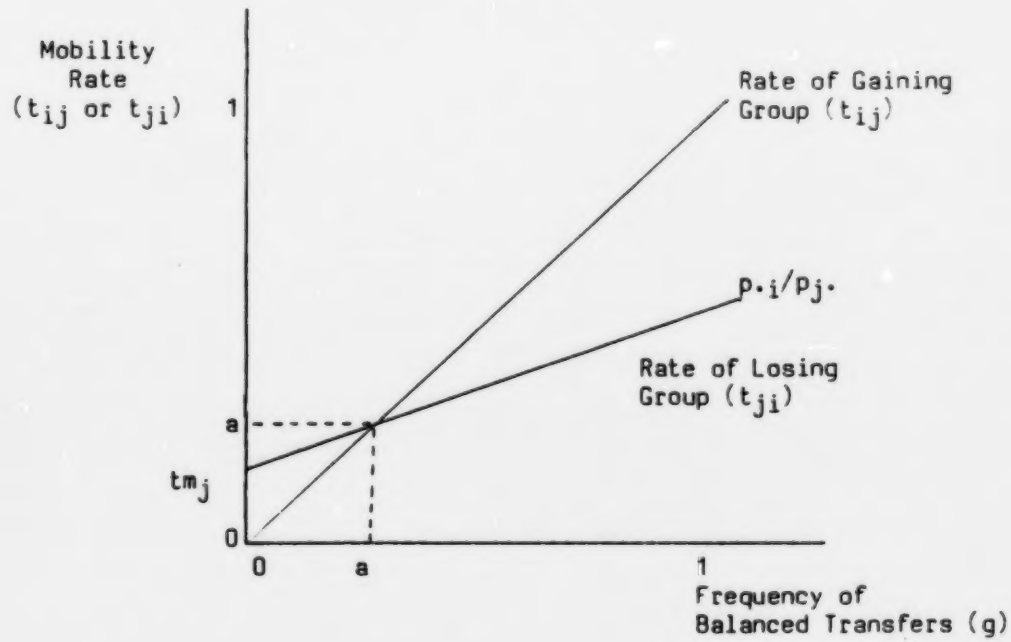
$$t_{ij} > tm_j + t_{ij}\, p_{i\cdot} / (1-p_{i\cdot})$$

$$t_{ij}\,(1-2p_{i\cdot}) > tm_j\,(1-p_{i\cdot})$$

$$t_{ij} > (p_{\cdot i}-p_{i\cdot}) / (1-2p_{i\cdot}) = a \qquad \text{(XXV)}$$

It can also be shown that $t_{ji} > a$. In other words, when $g > a$, we have $t_{ij} > t_{ji} > a$; similarly, when $g < a$, we have $a > t_{ji} > t_{ij}$. The mobility rate lines cross when g is equal to a (Figure 2).

Figure 2. The Gaining Group is in the Minority



Consequently, there is nothing paradoxical about the observed situation of the English and French groups in Quebec. Whenever the gaining group is in the minority, it is observed when g is larger than a.

To sum up, in many cases we know the marginal totals of a table, but not the values of their components. Examining the mother-tongue composition ($P_{i.}$ and $P_{j.}$) and the home-language composition ($P_{.i}$ and $P_{.j}$), we see that group i dominates group j ($P_{ji} - P_{ij} = P_{.i} - P_{i.} > 0$). Given this information, what can we say about the mobility rates of the two groups (t_{ij} and t_{ji})? We determined that the losing group's mobility rate (t_{ji}) is invariably higher than that of the gaining group (t_{ij}) when the latter is in the majority. The same situation occurs when the gaining group is in the minority, provided the two mobility rates are below a certain level (a), which depends solely on the table's marginal totals. Above this threshold the situation changes, and we find that the gaining group's mobility rate (t_{ij}) exceeds that of the losing group (t_{ji}).

In the course of our analysis, we discovered two fundamental parameters. The first, the losing group's net linguistic mobility rate (tm_j), measures the structural effect of language mobility; together with the proportion of the gaining group in the initial population ($P_{i.}$), it summarizes all the information contained in the margins of the table. The second parameter, g, measures the frequency of balanced language transfers. It carries all the information available in the body of the table, since a two-by-two table whose marginal totals are known has only one degree of freedom.

Let us take an example. According to the 1981 Census, there were 601,555 people in Quebec whose mother tongue and home language were both English; 82,135 had transferred from English to French; shifts from French to English numbered 106,370; and 5,128,135 people had French as mother tongue and home language. Hence, language transfers clearly favour the English group. The French group's net linguistic mobility rate (tm_j), or the net Anglicization rate, was +0.46 per cent. Yet the English group was found to have a much higher mobility rate (12.0 per cent) than the French group (2.0 per cent). This is because these rates were far above the threshold defined by a, which was very low in this

case (0.53 per cent). We shall denote the English group by the letter i and the French group by the letter j. The last equation in (XXIII) can be transformed into the following:

$$p_{ji} = (p_{.i} - p_{i.}) + p_{ij}$$

and $$n_{ji} = (n_{.i} - n_{i.}) + n_{ij} \qquad \text{(XXVI)}$$

The first term on the right side represents the structural effect of language transfers; balanced transfers are given by the second term. Thus, in our example we have

$$106{,}370 = 24{,}235 + 82{,}135$$

Changing the value of g, here 12.0 per cent, will obviously have no impact on the structural effect. So let us try replacing it with the value of a (0.53 per cent); we obtain

$$27{,}876 = 24{,}235 + 3{,}641$$

The proportion of balanced transfers drops from 77 per cent (82,135/106,370) to 13 per cent (3,641/27,876). Consequently, it would taken a very sharp decrease in the number of transfers to make the mobility rates equal (at 0.53 per cent), and an even sharper one to reverse their size ranking. The English-French language mobility situation in Quebec is therefore no surprise at all. Moreover, an analogous situation exists in the rest of Canada between the French group and the "other" group (Tables 1 and 4).

To solve this problem, Castonguay (1980) proposed a different approach from the one we have just described. It entails the assumption that the mobility rates are proportional to the weight of the final-language group in the initial population:

$$t_{ij} = b_{ij}\, p_{j.}$$

and $$t_{ji} = b_{ji}\, p_{i.} \qquad \text{(XXVII)}$$

It seems logical to convert the b_{ij} and b_{ji} into standardized mobility rates, as follows:

$$b_{ij} = p_{ij} / p_{i.}p_{j.}$$

and

$$b_{ji} = p_{ji} / p_{i.}p_{j.} \qquad \text{(XXVIII)}$$

These rates ignore the structural changes caused by language mobility. This means that there may be large differences in their ranges, which makes comparisons difficult.

The method we have developed can easily be extended to apply to any number (n) of languages. All that is required are three sets of independent parameters:

(1) mother-tongue composition (n-1 parameters);
(2) the net effect of transfers between each pair of languages (n(n-1)/2 coefficients);
(3) the intensity of balanced mobility between each pair of languages (n(n-1)/2 parameters).

Here are a few guidelines for the computation of non-trivial parameters. The first step is to determine which group gains and which loses in each pair of languages. We may assume, without loss of generality, that group i prevails over group j. Let $m_{ji} = p_{ji} - p_{ij}$ and $m_{ij} = 0$. This will enable us to calculate the minimum value of each mobility rate tm_{ij} (= $m_{ij}/p_{i.}$). For each pair of languages, the highest rate (there are n(n-1)/2 of them) measures the net effect of language transfers. It is somewhat more difficult to determine the value of the g parameters (which also number n(n-1)/2). We shall omit the details of the arguments, but it can easily be verified that

$$g_{ij} = \min (p_{ij},p_{ji}) / \min \left(p_{.i} - \sum_{l \neq i} m_{li},\ p_{.j} - \sum_{l \neq j} m_{lj}\right)$$

$$= \min (p_{ij},p_{ji}) / \min \left(p_{i.} - \sum_{l \neq i} m_{il},\ p_{j.} - \sum_{l \neq j} m_{jl}\right) \qquad \text{(XXIX)}$$

This implies that $g_{ij} = g_{ji}$, as might be expected.

These indexes reveal a clear distinction between two types of language mobility: (1) mobility that can be described as structural, in that it has an impact on the demolinguistic structure; and (2) exchange mobility, which measures the frequency of balanced language transfers. These are two extremely important aspects of language mobility, and the approach outlined here makes them accessible. Nevertheless, it is not the only available method of analysis.

From Mobility Rates to Linguistic Composition

When we know only the mobility rates (t_{ij} et t_{ji}) and the retention rates (t_{ii} et t_{jj}), what can we say about the mother-tongue composition and the home-language composition? In the absence of additional information, let us suppose that the population is in a state of equilibrium, which amounts to assuming that $p_{i.} = p_{.i}$ and that $p_{ji} = p_{ij}$. We shall denote this equilibrium composition by p^*_i and p^*_j ($= 1-p^*_i$). Accordingly, we have

$$p^*_i \, t_{ij} = (1-p^*_i) \, t_{ji}$$

$$p^*_i = t_{ji}/(t_{ij} + t_{ji})$$

$$p^*_j = t_{ij}/(t_{ij} + t_{ji}) \qquad \text{(XXX)}$$

These relations do not show which group gains from the real language exchanges. The only thing we can say is that if $t_{ji} \geq t_{ij}$, we have $p^*_i \geq p^*_j$, and conversely, that if $t_{ij} \geq t_{ji}$, we have $p^*_j \geq p^*_i$. In other words, group i makes up the majority of the equilibrium composition when its mobility rate is lower than that of the other group.

To see this more clearly, let us compare the equilibrium composition with the observed compositions. To simplify the work, we shall set

$$v = p_{ji} - p_{ij} = p_{.i} - p_{i.}$$

$$k = t_{ij} + t_{ji}$$

and $w = v/k$ (XXXI)

We can also write

$$v = (1-p_{i.})\, t_{ji} - p_{i.}\, t_{ij}$$

Replacing the mobility rates with values derived from (XXX), we obtain

$$v = (1-p_{i.})\, p_i^*\, k - p_{i.}\, (1-p_i^*)\, k$$

$$v = k\, (p_i^* - p_{i.})$$

and finally $w = p_i^* - p_{i.}$ (XXXII)

Since k is greater than or equal to zero, v must have the same sign as w. This implies that if i is the gaining group, its weight in the equilibrium composition is larger than in the mother-tongue composition. The opposite is true when i denotes the losing group.

Net we shall compare the equilibrium composition with the home-language composition. Letting $u = p^* - p_{.i}$ et remarquons que:

$$p_i^* - p_{.i} = (p_i^* - p_{i.}) - (p_{.i} - p_{i.})$$

$$u = w - v$$

$$= w\,(1-k) \qquad \text{(XXXIII)}$$

If we assume that k ($= t_{ij} + t_{ji}$) is less than 1, it follows that u must have the same sign as w. Except in highly unusual circumstances ($k > 1$) not likely to arise in the study of language mobility, there are ultimately only three possibilities to consider: (a) group i gains ($P^*_i > P_{\cdot i} > P_{i\cdot}$); (b) the two groups are in a state of equilibrium ($P^*_i = P_{i\cdot} = P_{\cdot i}$); and (c) group i loses ($P_{i\cdot} > P_{\cdot i} > P^*_i$).

When the gaining group (i) is in the minority, its mobility rate (t_{ij}) may of course be higher than that of the losing group (t_{ji}). On the basis of previous results, however, we can say that the following must be true:

$$t_{ji}/(t_{ij} + t_{ji}) = P^*_i > P_{\cdot i} > P_{i\cdot} \qquad \text{(XXXIV)}$$

Thus, we find that the equilibrium linguistic composition is simply a transformation of the mobility rates.

It is fairly easy to generalize this method for any number of languages. This can be done using the theory of finite Markov chains (Bernard, 1978; Boudon, 1973; Gordon, 1965; Kemeny, Snell and Thompson, 1965). This model can be presented in two ways. Suppose we have a table of language exchanges between n languages. With these data, we calculate $n(n-1)$ mobility rates and n retention rates (the latter can be deduced from the former). The theory of stationary Markov chains implies that there is a state of stable equilibrium such that mother-tongue composition is identical with home-language composition. The equilibrium composition depends exclusively on the mobility rates. By comparing it with the observed compositions, we can often separate the gaining groups from the losers. This is the method we have been using so far.

In the study of language mobility in Canada, three classes of languages are usually considered: English, French and a residual class, the other languages. As an illustration, we shall examine the discrepancy between the equilibrium composition and the linguistic compositions observed in Quebec in 1971 in the 25-44 age group:

	Mother Tongue	Home Language	Equilibrium
English	12.1	14.3	19.0
French	80.1	80.3	79.9
Other	7.8	5.4	1.1
Total	100.0	100.0	100.0

The process works to the English group's advantage: 12 per cent of the population reported English as mother tongue and 14 per cent as home language; in the equilibrium composition, the proportion rises to 19 per cent. The converse is true for the "other" languages, which just barely top the one-per-cent mark in the equilibrium state. The French group's position is more complex. It is a situation that cannot arise when only two languages are involved. On the basis of the observed compositions, French benefits marginally from the process since it is reported more frequently as a home language (80.3 per cent) than as a mother tongue (80.1 per cent); its gains from the "other" languages slightly outweigh its losses to English. In the state of equlibrium, however, its proportion (79.9 per cent) is slightly smaller than its mother-tongue percentage. This is due to the fact that the "other" languages make up only a tiny part of the equilibrium composition.

If in addition to mother tongue and home language, we knew the mother tongue of the mother (or the father), it would be possible to compute a new state of equilibrium encompassing all three linguistic compositions. Can ethnic origin be considered a satisfactory approximation to this variable? Let us make that assumption. In this new equilibrium composition, the English proportion increases further (19.7 per cent), as does the French proportion (80.3 per cent), whereas the "other" languages virtually disappear. This has to do with the fact that in the populations of British and French origins, the mobility rates from English and French to the "other" languages are practically zero. In sum, the French group just manages to stay level, and the net effect is the same as if all members of the "other" group were destined to make English their main language.

There is another, more conventional interpretation of the equilibrium composition. It is founded on an implicit model of population change in language groups. Suppose we have a population that is closed to migration and has uniform mortality and fertility. Let us further assume that the language mobility rates apply to women of childbearing age and that children's mother tongue is the same as their mother's home language. Finally, we shall assume that the process is stationary, that is that the mobility rates do not fluctuate over time. In this case, the mother-tongue composition of the children (first generation) will be equal to the home- language composition of their mothers (initial generation). If we then apply the same mobility rates to the first generation, we will obviously obtain their home-language composition; this in turn will be the mother-tongue composition of the second generation, and so on. This process traces the effects of the mobility rates on the linguistic compositions from generation to generation. It tends to a limit composition that is independent of the initial state, depending solely on the mobility rates.

Clearly, the limit composition merges with the equilibrium composition. Yet it would seem that the model on which the equilibrium composition is based is more realistic than the model that generates the limit composition. The latter's predictive behaviour is suspect. To avoid any confusion, it is best to stick to the equilibrium composition concept.

Nevertheless, it is sometimes helpful to relate a process to a specific point in time. Let us attempt to determine the mother-tongue composition of the first generation (generation 1), assuming that fertility (D) in the initial generation (generation 0) varies by home language. Thus we have

$$n_{i.}^{(1)} = \sum_j n_{ji}^{(0)} D_i = D_i \, n_{.i}^{(0)}$$

$$\text{and} \quad n_{..}^{(1)} = D \, n_{..}^{(0)} \qquad \text{(XXXV)}$$

Average fertility (D) can be expressed in terms of the fertility of each language group:

$$D = \sum_i p_{.i}^{(0)} D_i \qquad \text{(XXXVI)}$$

The first generation's mother-tongue composition depends on both language mobility and differences in fertility:

$$p_{i.}^{(1)} = p_{.i}^{(0)} D_i/D \qquad \text{(XXXVII)}$$

The process is in the equilibrium state when mother-tongue composition does not vary from one generation to the next. For this, it is sufficient that:

$$p_{i.}^{(0)} = p_{.i}^{(0)} D_i/D \qquad \text{(XXXVIII)}$$

This condition is equivalent to supposing that the ratio of the average fertility (D) to the fertility of the language group being studies (D_i) is equal to the group's continuity index (c_i). In this way, the effects of language mobility can be interpreted in terms of differential fertility (Lachapelle and Henripin, 1982).

Under current mortality conditions in Canada, the population will replace itself if women have an average of 2.1 children. Let us set D = 2.1 and, purely for illustrative purposes, calculate (on the basis of Table 3) the completed fertility levels required to counterbalance the effects of linguistic mobility:

	Canada	Quebec	Canada excl. Quebec
English	1.9	1.8	1.9
French	2.2	2.1	2.9
Other	3.8	2.9	4.0
Total	2.1	2.1	2.1

The Anglophone group could get by with a fertility level below the replacement level since it gains from language transfers. For Francophones in Quebec, the situation is the same as if there were no language mobility; their fertility level must be equal to the replacement level. Francophones and allophones outside Quebec require, assuming zero migration, completed fertility levels of roughly three or four children per woman just to replace their poulations. These are, of course, only approximations.[4]

Overview

The initial phase in the analysis of a phenomenon is often characterized by a wide variety of approaches and the production of many indexes. In this way, the various facets of the phenomenon are explored and its principal trends uncovered. Occasionally, however, this process produces apparently contradictory findings, which creates the impression of confusion and casts doubt on all the conclusions. In Quebec, the balance of language transfers between the English and French groups benefits the former at the expense of the latter, which is hardly surprising; yet the English group's language mobility rate is consistently much higher than that of the French group. This is a paradox. Net transfers or mobility rates, which one provides a more accurate picture of the language situation?

Without an in-depth methodological study, the choice is very likely no more than a subjective preference. To go beyond this stage, it is necessary to carry out a critical review of the analytical methods and indexes used. This is not a simple task. One must first go to the heart of the matter and focus on what may be referred to as the fundamental table. It shows the joint distribution of a given population by mother tongue and home language.

For the simplest case, in which only two languages are involved, the table has four degrees of freedom. If the absolute frequencies are converted to relative frequencies, so that the grand total is unity, there will be only three degrees of freedom. This means that there can be only three independent coefficients. Two approaches to selecting them were proposed.

The margins of the fundamental table have at most two degrees of freedom. We have all the information they contain when we know the weight of the dominant group in the mother-tongue composition and a coefficient expressing the structural imbalance caused by language mobility. A number of indexes can be used to measure this imbalance. We selected

the net linguistic mobility rate of the losing group. When these two parameters are held fixed, there is only one degree of freedom left in the body of the table. It is possible to define a parameter that expresses the intensity of the balanced linguistic mobility in a population. All other rates or indexes can be expressed in terms of the three coefficients selected. This approach can readily be extended to apply to any number of languages. In particular, it demonstrates that there is nothing paradoxical about the language situation in Quebec.

The second approach consists in comparing the two mobility rates with the observed linguistic compositions. It can easily be shown that all the information that the table holds is contained in the two mobility rates and the proportion of one of the groups in the mother-tongue composition. Yet how can rates be compared with compositions? It was proved that the mobility rates can be transformed into what was termed an equilibrium composition. This transformation is nothing new; it is a conventional application of the theory of stationary Markov chains.

The first approach draws a highly relevant distinction between the structural effect of linguistic mobility on one hand and exchange mobility on the other. Over the past decade, the spotlight has been trained on the former. It might be productive to consider both aspects of the phenomenon. Although the first approach seems better suited to the study of language mobility than the second, the latter may be very useful for special problems.

Notes:

1. In Canadian censuses, mother tongue does not refer to the first language spoken in childhood but rather to the first language learned (or spoken) and still understood.

2. Home language is defined as the language spoken most often at home.

3. It can readily be shown that this situation can be combined with the case where the dominant group is in the majority. The mobility rates have the same range.

4. This interpretation of the continuity index was discussed at length in Lachapelle and Henripin, 1982.

References

BERNARD, Jean-Thomas. (1978), "La mobilité linguistique et l'attraction des langues au Québec". (Quebec, Université Laval, Economic Department) (multigr).

BOUDON, Raymond. (1973), Mathematical Structures of Social Mobility, (San Francisco: Jossey-Bass), 168 p.

CASTONGUAY, Charles. (1980), "Sur quelques indices de propension à l'exogamie et au transfert linguistique". Cahiers québécois de démographie: 9, 3, 53-70.

CASTONGUAY, Charles. (1983a), "Exogamie et transferts linguistiques chez les populations de langue maternelle française au Canada". Démographie et destin des sous-populations. Colloque de Liège, September 21-23, 1981), (Paris, Association internationale des démographes de langue française): 209-215.

CASTONGUAY, Charles. (1983b), "Évolution des transferts linguistiques au Québec selon les recensements de 1971 et de 1981", Québec, (Conseil de la langue française, temporary edition): 76 p. (multigr.).

CASTONGUAY, Charles and MARION, Jacques. (1974), "L'anglicisation au Canada", Bulletin de l'Association des démographes du Québec: 3,1, 19-40.

DE VRIES, John and VALLEE, Frank G. (1980), Usage de la langue au Canada. (Ottawa: Ministry of Supply and Services Canada): 187 p.

GORDON, Patrick. (1965), Théorie des chaînes de Markov finies et ses applications. (Paris, Dunod): 132 p.

GRYZ, Zbigniew Jan. (1980), Socio-Demographic Determinants of Language Shifts in Canada. (Ph. D. Thesis, Ottawa, Carleton University, Dept. of Sociology and Anthropology): xxxi + 411 pages.

KEMENY, J.G., SNELL, J.L. and THOMPSON, G.L. (1965), Algèbre moderne et activités humaines. (Paris, Dunod): 344 p.

KRALT, John. (1977), "The Case Against Net-Language Transfer: An Alternate Measure", in LAMY, Paul, Language Maintenance and Language Shift in Canada. (Ottawa, University of Ottawa Press): 75-79.

LACHAPELLE, Réjean. (1980), "Evolution of Ethnic and Linguistic Composition", in BRETON, R., REITZ, J.G., and VALENTINE, V.F., Cultural Boundaries and the Cohesion of Canada, (Montreal, The Institute for Research on Public Policy): 15-43.

LACHAPELLE, Réjean. (1983), "Définition et analyse des mobilités démographiques: l'exemple de la mobilité linguistique". Démographie et destin des sous-populations. (Colloque de Liège, September 21-23, 1981), (Paris, Association internationale des démographes de langue française): 237-248.

LACHAPELLE, Réjean and HENRIPIN, Jacques. (1980), La situation démolinguistique au Canada: évolution passée et prospective. (Montreal, L'Institut de recherches politiques): xxxii + 391 p.

ROCHON-LESAGE, Madeleine and MAHEU, Robert. (1974), "Composition ethnique et linguistique de la population du Québec". Annuaire du Québec. (Quebec, Éditeur officiel du Québec): 206-212.

SOCIAL AND ECONOMIC STUDIES DIVISION

RESEARCH STUDIES SERIES

No.

1a. Transforming Monthly Uniform Crime Reports to Incident Records
Craig McKie, revision, Summer 1983.

1b. Transformation des données du programme de déclaration uniforme de la criminalité en fonction des données sur les cas
Craig McKie, révision, été 1983

2. The Use of "Fixed Coefficient" Matrices to Relate Occupation and Field of Study to Manpower Requirement Models
Garnett Picot, Spring 1982

3. Inconsistencies in the Occupational Coding of the Same Individual in Two Surveys
John Turner, Spring 1983

4. Sizing Up the Burden of Unemployment in Canada
Part I - A General Labour Market Perspective
R. Paul Shaw, Spring 1984

A revised version of this paper has been published in Canadian Public Policy, Vol. XI, no. 2, June 1985, 143-160.

5. Sizing Up the Burden of Unemployment
Part II - A Perspective on Financial Hardship
R. Paul Shaw, Summer 1984

6. La ruralité québécoise: Évolutions et perspectives
Clermont Dugas, été 1984

A revised version of this paper has been published in Cahiers de géographie du Québec, vol. 28, nos 73-74, avril-septembre 1984, 183-204.

7. Fiscal vs Traditional Market Variables in Intermetropolitan Migration in Canada 1956-1981
R. Paul Shaw, Summer 1984

8. Historical Statistics on Justice
Paul Reed, Summer 1984

9a. Life Table Methodology
Canada and Provinces (1980-1982 cycle)
Dhruva Nagnur, Fall 1984

9b. Méthode de calcul des tables de mortalité
Canada et provinces (cycle 1980-1982)
Dhruva Nagnur, automne 1984

10a. Heterogeneity and Selection in Population Analysis
Nathan Keyfitz, Fall 1984

10b. Hétérogénéité et effet de sélection dans l'analyse des populations
Nathan Keyfitz, automne 1984

11. The Industrial Mix of Employment: How and Why it Changes A Review of the Past Three Decades
W. Garnett Picot, Spring 1985

12a. Analyse de la mobilité linguistique: indices, observations et modèles
Réjean Lachapelle, printemps 1985

12b. Analysis of Language Mobility: Indexes, Observations and Models
Réjean Lachapelle, Spring 1985

13a. An Analysis of Expenditure Patterns of the Elderly: Cohorts - Going Through Life Together
Louise Heslop, Spring 1985

13b. Analyse des dépenses des personnes âgées: cohorte, âge et période
Louise Heslop, printemps 1985

14a. Note de recherche - La composition linguistique: réalité et perception
Réjean Lachapelle, printemps 1985

14b. Research Note - Linguistic Composition: Perception and Reality
Réjean Lachapelle, Spring 1985

15a. Rectangularization of the Survival Curve and Entropy The Canadian Experience
Dhruva Nagnur, Summer 1985

16a. Sizing up the Burden of Unemployment in Canada Part III: Impact of Unemployment on Family Income
Paul Shaw, Summer 1985